Laying Tracks as the Train Approaches

INNOVATIVE OPEN ACCESS BOOK PUBLISHING AT HEIDELBERG UNIVERSITY FROM THE EDITORS' POINT OF VIEW

ANDREA HACKER and ELIZABETH CORRAO

In April 2016, Heidelberg University's newly founded open access publisher heiUP launched the first volume of the new book series Heidelberg Studies in Transculturality. This article reports on the challenges, accomplishments, and setbacks that informed the entire editorial production process, not only of the first volume but also of the series and the publishing enterprise overall. The authors offer insights on crucial issues that any new open access publishing endeavour at an institution might face, namely acquiring manuscripts, designing and building workflows, and collaborating with partners to build an outlet for hosting the finished product. This article also illustrates how the goal of providing a new digital reading experience through an innovative HTML format, in addition to print-on-demand and PDF versions of each manuscript, affected the progress of the entire project. Finally, we report on what it took to deliver results.

Keywords: open access, XML, editing, academic publishing, university press

INTRODUCTION

On 29 April 2016, the first volume in a new open access gold series called Heidelberg Studies in Transculturality went live on the newly founded Heidelberg University Publishing (heiUP) website.[1] Both the series and heiUP are the result of a joint project of Heidelberg University's Cluster of Excellence 'Asia and Europe in a Global Context: The Dynamics of Transculturality'[2] (specifically its publications office) and Heidelberg University Library. The project is funded with a two-year grant from the German Research Council (DFG) as part of its funding program called 'scientific monographs and book series in open access' (wissenschaftliche Monographien und monographische Serien im Open Access).[3]

This article reports on three major editorial objectives that the authors of this essay have worked toward, as well as progress made so far: 1) to acquire and edit quality manuscripts in a very young academic discipline; 2) to develop a production workflow for a new English-language book series; and 3) to collaborate simultaneously with the university library and the university rectorate to build a university press for the finished products.

We approached these objectives with the benefit of several years of publishing expertise at the Cluster, where, in 2010, Andrea Hacker, Cluster's managing editor, began to build the workflows for an open access gold ejournal, *Transcultural Studies*, and an earlier book series published in collaboration with Springer.[4] These workflows covered managing original book manuscripts from submission to fair copy with Springer, and managing the entire production of the e-journal from submission to publication and distribution. The combined expertise eventually informed the new book series—and for that matter much of heiUP.

Parallel to these developments, the university library significantly increased its open access activities with the establishment of an open access green repository, retro-digitization projects, an open access fund, and hosting a quickly increasing number of ejournals on Open Journal Systems (OJS).

In autumn 2012, when the DFG published the call for applications for the aforementioned open access program, the university library and the Cluster joined forces. Andrea Hacker and the Cluster's former senior software engineer, Dulip Withanage, co-wrote an application with their university library colleague Martin Nissen under the auspices of Cluster director Axel Michaels and the university library's director Veit Probst.

While the team awaited the DFG's decision, the university's rectorate decided independently to pursue the establishment of Heidelberg's own university press. The idea had been circulating for years, but now that technological infrastructure and concrete publication structures were taking shape within the institution, the university leadership felt it was the right moment to move ahead. The rectorate convened an advisory board, which comprised representatives from across campus, to discuss strategies and first steps.

The two projects quickly merged. In December 2013, the DFG approved the application for the pilot series and the two-year funding made it

possible to hire a designer and technical editor in the library and, in the Cluster, a full-time programmer and a dedicated half-time editorial assistant. Elizabeth Corrao, co-author of this essay, filled the latter position in May 2014. Her task was to assist in various aspects of press development and the editorial workflow, including copywriting and editing, developing the style guide, and collaborating with the library staff to create content for the press website.

The university leadership's strategic goal of testing the feasibility of a digital publishing venture dovetailed with the DFG project's remit to develop a business model for the production of open access monographs; the positions were quickly filled and the project was underway.

It soon became obvious that countless aspects and tasks apart from manuscript development would have to be considered and managed when building a new book series alongside the university press that would eventually publish it. These ranged from designing a bilingual website and book covers that adhered to university-approved branding to clearing the tax status of the young undertaking. To tackle these aspects, a working group was convened to coordinate and align the university's goal of launching an open access publishing venture with the pilot project. Progress was reported to heiUP's advisory board.

After the first few meetings, the working group realized that the development of a business model would have to be preceded by an experimental phase in which many technical—as well as managerial—aspects of the press would have to be developed, improvised, tested, and stabilized. A reliable analysis of production costs for the series, or even just one book, could only be made once a relatively stable workflow was in place. The following will report on the building of this workflow from the editorial team's point of view.

A New Kind of Book: The Added Value of an Innovative Format

The developments at Heidelberg University took shape against a wider background of rapid innovation in digital publishing, and open access publishing in particular. At first the resulting changes were observable primarily in the realms of repositories and journals, but they gradually emerged in book publishing as well. The question of how the scholarly monograph could be reimagined was therefore central to the project from its inception.

Exploring and expanding the digital potential of the book in terms of workflow—of which more below—and in terms of genre had to go hand in hand with any attempts to determine how the production of such reinvented books could become sustainable. The idea was to nudge the experience of reading books in the humanities and social sciences beyond the usual technological boundaries. Our project intended to offer something in addition to the printed book, which is still so crucial to scholars' careers, and the PDF, which by now has, for better or for worse, been established as a standard file format for articles and books alike.

While it was clear from the start that our books had to appear in both these formats—a print-on-demand (PoD) solution was as much a sine qua non as the PDF—it became increasingly apparent that an HTML version would be flexible enough to offer room for innovative presentation. An early fan of eLife's open-source viewer Lens (previously eLENS),[5] Andrea Hacker insisted—at times in the face of concerns over development challenges—that this added value would give our young undertaking the unique chance to present humanities and social science research in an innovative way to an audience that increasingly reads research material online. A viewer like the eLife Lens enables publishers to move away from the linear reading experience of the PDF and toward a dual-panel layered experience where the flow text appears on one side of the screen accompanied by supplementary information—references, illustrations, multimedia, and text navigation—on the other. The benefit of such a split-screen viewing format is that the various layers of concomitant information that most humanities and social science texts rely on become disentangled and can be viewed individually, without losing their connection to the flow text.

A final consideration was the fact that, unlike the linear book format, a layered HTML presentation not only enhances the content with traditional multimedia source material such as film, audio, and images; it can also make way for meaningfully incorporating other kinds of research results that are fast becoming alternative ways of performing scholarly communication, such as digital humanities projects, databases, or software.[6]

THE PRODUCT AND THE PROCESSES REQUIRED TO REALIZE IT

To manage the logistical tasks associated with producing a traditional PDF as well as an innovative HTML format, we faced a twofold challenge: recruiting new authors and inventing a production workflow.

Winning Authors for an Innovative Book Series with an Unproven Publisher

One of the greatest challenges for open access and scholarly publishing overall is the acquisition of high-quality content.[7] This may seem counterintuitive considering the current information deluge and inflationary increase of publications, but it is something that all publishers struggle with: while the pressure to publish or perish remains ubiquitous in academia, there are few authors who commit important work (in most cases for free) to established presses, and even fewer to brand-new open access outfits, no matter how well meaning they may be. After all, authors are keen to find a prestigious publisher for their research. For our book series, acquisition is and will remain for the foreseeable future the primary challenge: the series editors and the managing editor have to convince authors that their content is in good hands, despite the series being new, the academic field of transcultural studies itself not being much older, and the nascent publisher, heiUP, not (yet) offering much in the way of prestige.

Andrea Hacker followed a three-part strategy to obtain manuscripts. First, she promised authors a publication that featured the full spectrum of currently possible reading experiences: PoD (traditional), PDF ('tradigital'), and an innovative form of HTML. Second, she committed to offering our first few authors and book editors extra support with their manuscript development, including assistance with rewrites and language editing. Finally, she showed prospective authors the potential reach of publications in these formats and argued for the overall good cause of sharing knowledge openly. She supported all these arguments and sparked authors' imaginations by showcasing the technological possibilities of innovative HTML display with the eLife Lens, along with examples from our ejournal.

Since the series was always intended to function as a publishing instrument for transcultural research, a young academic discipline in the humanities and social sciences that has one of its strongholds in Heidelberg, this first round of acquisition took place among researchers at the Cluster: visiting scholars, post-docs, and tenured professors alike.

The results were surprisingly positive, particularly when compared to the resistance encountered some five years earlier when we began acquiring essays for our open access ejournal. Authors' overall willingness to experiment, as well as their acceptance of open access publishing in

general, was truly encouraging.[8] We currently have ten book projects slotted for the series: one has been published, one is in production, one is in rewrite, one is in peer review, and the others are being written.

Developing a Production Workflow while Building an Outlet for Our Books

This reassuring level of interest in our series immediately triggered another challenge: once we obtained confirmation from authors that we would receive their manuscripts, we were faced with the task of developing an entire editorial and production workflow to support these manuscripts from submission to publication and beyond. We found ourselves in the proverbial position of laying tracks while the train was already barreling toward us.

Although we could build on previous editing experience, we still needed myriad tools and steps for the workflows that underlie the actual publishing process, ranging from a comprehensive style guide for authors to peer-review standards. In terms of production, we had to develop and agree not only on how best to prepare manuscripts for an as-of-yet improvised workflow but also on the design of the end result, including layout, fonts, covers, and eLife Lens customization. The two central areas that affected us the most as editors, and that had to be created simultaneously, were the editorial policies and editing technology.

The first step in building the editorial process was to convene an editorial board for the new series. Since transcultural research more often than not works across academic fields, we needed experts who could assess and support a wide variety of disciplines as well as judge the soundness of the approaches taken to reflect the transcultural dimension of submitted manuscripts. After some internal consultation, the Cluster leadership produced a list of ideal candidates, all of whom we were able to convince to join our undertaking.[9]

The editorial board's first task was to discuss and ratify the series' editorial policy. We wrote an initial draft that combined the pre-existing policies from our ejournal and Springer series with relevant details we gleaned from the policies and procedures sections of other open access publishers' websites. This draft was sent to the editorial board for input, adjusted according to the feedback, and subsequently fine-tuned. The result reflects a comprehensive vision for our series and covers questions ranging from author licensing choices (CC-BY to CC-BY-NC-ND) to plagiarism and peer-review policies.

The next step was to design an editorial workflow for submitted manuscripts. Here, too, we were able to draw on our previous publishing experience with the ejournal and the Springer book series. We adapted these workflows, discussed them with the series' editorial board, and consulted heiUP's advisory board for approval. Both the editorial policy and the editorial workflow from the ejournal served as the blueprint for heiUP's policies and workflows, covering issues ranging from quality control to information for authors.

Parallel to the development of our editorial policies and workflow, we had to master the technology necessary for our editing tasks while also attempting to develop it further to suit our purposes. We originally intended to use Open Monograph Press (OMP), a new open-source publishing platform developed by the Public Knowledge Project, the same organization whose OJS software we have used for the ejournal for a number of years and with whom we had developed a close working relationship. However, the platform's newness presented additional challenges in terms of ongoing technical developments and the availability of support, which meant we were not able to launch the press with our intended workflow and instead had to improvise with a decentralized manuscript management system using email, Microsoft SharePoint, and local servers.

Adding to this challenge was our ambition to create an easily usable, single-source publishing workflow (Word to XML to HTML and PDF) that would eliminate the separate Word-to-XML-to-HTML and Word-to-XML-to-PDF conversion, which consumes a lot of time and is also prone to error. The plan had been to build a 'what you see is what you get' (WYSIWYG) editor that would allow users unfamiliar with XML coding to produce reliable XML, which could then, with the help of style sheets, be rendered automatically into HTML and PDF. The central technical concern of our project has always been to create a complete open-source workflow centred on a single-source, clean XML that would enable editors to produce a variety of end formats from text files generated in programs such as MS Word or OpenOffice.

OBSTACLES AND SETBACKS

In addition to the challenges associated with acquiring book manuscripts and developing the editorial workflow, while also learning and developing the technology needed for editing, the project encountered various obstacles and setbacks along the way.

Task Shifting after Personnel Turnover

One of the most challenging of these setbacks was personnel turnover. For example, the coder originally hired to develop the WYSIWYG editor left the project for personal reasons less than six months into her contract, causing the development of the WYSIWYG and template to come to a sudden halt. While our colleagues at the library eventually took over various aspects of this role, the lack of a dedicated staff person caused serious delays in our initial production timeline of two years. While waiting for an operational WYSIWYG editor, we have followed a 'plan B': the PDFs are created in Adobe InDesign based on Word documents, while the HTML is drawn from an XML file that is initially generated from Word documents through meTypeset, which then has to be manually adjusted.[10] Adopting this workflow meant that, in order to meet production deadlines, Elizabeth Corrao had to take on technical responsibilities beyond her role as editor. These roles, which would normally fall to a production team, included running the typesetter to check the quality of XML and editing the XML itself to resolve numerous formatting issues that arose when we tried to automatically integrate external programs such as Zotero into the text.

This is just one example of how an ad hoc publishing project can be forced to adjust and expand to rapidly changing circumstances in order to succeed. Editors become quasi-coders and programmers become unplanned inductees into the complexities of formatting and bibliographic citation style.[11] At the time of this article's writing, production still follows this plan B, although heiUP has been able to fill the vacant junior coding position with a programmer who is currently assisting with the adaptation of OMP. The development of the WYSIWYG editor will be resumed in the future.

Coping with Time Pressure and Facing Compromise

Our second greatest obstacle was time pressure and its ramifications for the overall goal of developing a stable enterprise that could eventually yield a feasible business model for the production of open access gold books. Because our first choice of technology remained out of reach and left us with a complicated workaround, the production process grew lengthy and was frustratingly slow at times. Consequently, we found ourselves negotiating between striving for timeliness (in terms of meeting our production goals) and volume (in terms of our hoped-for output

over the course of the first two years) while not wanting to settle in terms of quality. However, deciding where to focus our energies became increasingly challenging. The simple fact is that producing workflows and technology takes a significant amount of time, a luxury we did not have, given the project's limited funding period, the editorial team's exceptionally small size (i.e., the authors of this article), and mounting pressure from within the institution to deliver results. A decision had to be made whether to cut corners to mimic a sustainable workflow, or to go all out to ensure that our series began with a strong volume that would also offer the open access publishing community something genuinely new. We felt that the book and the series that it spearheaded had to live up to the expectation of 'excellence' that the university—and with it the Cluster—carries in its name, so we opted for the latter and decided that the two of us would invest whatever it took to get it right.

The result was that the opening volume received a significant amount of editorial input and attention. We worked very closely with the author, Marie Sander, an anthropologist and alumna of Heidelberg University, who chose to turn her dissertation into a monograph for our series. While the cost of this development was far above and beyond what we could reasonably offer on a regular basis, we were able to get a sense of the amount of time that was really required to produce a carefully edited, well-designed, and glitch-free scholarly monograph.

Overall, the editing process consisted of three rounds of revision, throughout which we worked extensively with Marie Sander on each version of her 292-page text, exchanging approximately 250 comments using MS Word's track changes and over 300 emails. This process lasted from 4 May 2015 to 5 September 2015. During that time, the three of us held two conference calls of approximately two hours each, and Andrea Hacker held weekly individual calls with the author to discuss rewrites that lasted from two to three hours. After copy-editing was complete, nearly forty more hours were invested in formatting and standardizing the entire text because we still lacked a template that could effectively translate across all of our desired end formats. Once our colleagues in the library produced this template, a further three months of collaborative testing to identify and resolve glitches was required before our XML file was clean enough to be run through the typesetter and subsequently turned into HTML. From December 2015 to February 2016, Elizabeth Corrao proofread the entire text while running parallel tests with the

typesetter and correcting the monograph's Zotero bibliography for errors, completeness, and consistency. The XML-to-HTML conversion process was also quite extensive and fraught with significant setbacks, which meant that from December 2015 to April 2016, she met with the project's senior programmer, Dulip Withanage, approximately ten times, for a total of twenty-five hours, and exchanged approximately forty-five emails.[12] Once a clean XML and usable HTML had been created, she checked the final format and appearance of the HTML test files and corrected issues identified in the final HTML in the source XML. This process was completed last, in late April 2016. The softcover PoD proof was received at that time as well and approved on 25 April 2016.

Apart from these technical difficulties and the aforementioned shortcomings within the publishing workflow, two further circumstances protracted the time invested in producing our first book. First came the challenge of language and rewriting. Although Marie Sander knew the style of our prior book series and was already familiar with our other publishing endeavours, she is, like most of our current and prospective authors, a non-native English speaker. It is impossible to overstate the significant linguistic leap required of authors in the humanities and social sciences who write in a language other than their native tongue. This difficulty is compounded when their research focuses on a culture with a third language, which in Marie Sander's case was Chinese. To produce excellent books written by such authors, a publisher has to consider how to assure quality not only in the soundness of research — review mechanisms ought to take care of that — but also in language and structure. Finally, many junior researchers and first-time monograph authors struggle to refashion their dissertations into internationally competitive monographs. Ignoring the additional editing challenge this poses for the publisher is risky: peer review notwithstanding, transforming a dissertation into a book requires a significant amount of structural revision, which may need the helping hand of an editor for the monograph to reach its full potential.

CONCLUSION

Having reached the milestone of publishing the first book in our new series — in a likewise new bilingual publishing venture that we also helped to build — affords a critical look back at the overall development of the project. We began with a solid plan for a pilot project based on years of

experience; we were able to build on established networks in the open access community and utilize pre-existing publishing tools; we received generous funding for the project and met with the goodwill and support of the university leadership in what can be termed a historical chance to establish an innovative, open access publisher at one of Europe's oldest universities. Yet despite all of these auspicious conditions, the project had to overcome adverse circumstances that no one could have foreseen or controlled. While these included, as described above, significant technical setbacks, others were simply due to human nature, such as conflicting political goals, communicative short-circuiting, changing personal circumstances, or reassignments. They can all contribute significantly to the cost of getting an undertaking such as this off the ground.

The biggest setback of all, however, undoubtedly befell the plan to create a reliable business model based on the pilot project. The original plan had been to build a book series that would be financed by a mix of third-party funding, internal funding, some revenue from a 'freemium' model (PoD and one of the digital formats for sale, another freely accessible), and a low author processing charge to cover author copies and peer-review remuneration. The internal funding was to take the form of student assistantships for the copy-editing process, shifting some responsibilities of existing positions to allow for the development of operations, and creating, if possible, new positions dedicated to the increasing workload. Some of these ideas have already come to fruition—for example, the university library has hired new staff and restructured its management to accommodate the running of heiUP—while others have to be rethought or still pursued, due to the obstacles outlined in this essay. The fact remains that until we have a stable workflow and enough experience to optimize it, an exact cost analysis, and thus a reusable business model, stays out of reach.

The lesson learned is quickly summarized and self-evident: an ambitious project with so many goals, agents, and interests involved needs to budget and prepare for delays and difficulties, not only with time and resources but also with managerial flexibility so the project can quickly adapt to changing parameters.

We have also learned that there are certain pillars in a new publishing enterprise that can withstand the challenges of unforeseen change: a clear understanding of available resources, a realistic time frame, a collaborative work ethos, a manageable number of benchmarks, and unequivocally formulated and principled editorial goals are all necessary. Finally,

we hold fast to our conviction that while many processes in publishing can and should be automated, manuscript development is not one of them. Once the main instruments of production — from publishing platform to XML conversion and beyond — are available, more attention can be directed at the real concern of publishing academic research: ensuring quality.

Regardless of setbacks, which are part and parcel of any story, we are proud to have taken an important step toward offering an alternative, open access gold publishing model in the humanities and social sciences, by creating a bilingual open access gold publisher with a new, two-panel HTML format for books. There is, however, no time to rest on laurels; other important technological challenges are fast approaching. Some are issues of necessity to make the publishing venture functional, such as a conversion of LaTeX manuscripts to XML, reliable integration of open-source bibliographical tools like Zotero, or finalizing a fail-proof template for authors that will alleviate potential problems in rendering XML to our end formats. Other issues pertain more to experimentation and will have to be explored offering open peer review, integrating data sets and software projects into flow texts, and incorporating other network tools — from ORCID IDs to altmetrics — that have appeared in recent years. In short, we are far from being done. The train is still fast approaching, and we will continue laying tracks ahead of it.

ANDREA HACKER joined Heidelberg University's Cluster of Excellence 'Asia and Europe in a Global Context' as managing editor in 2009. She has since built several publishing and English for Academic Purposes (EAP) projects. She also publishes about open access and changes in scholarly communication.

ELIZABETH CORRAO joined Heidelberg University's Cluster of Excellence 'Asia and Europe in a Global Context' in 2013 as an academic writing consultant and graduate student assistant for the publications office. In 2014, she became the editorial assistant for the DFG-funded pilot project that this article outlines. In September 2016, she left the Cluster and joined the editorial team at EMBO Press, an editorially independent publishing platform for the development of EMBO scientific publications.

NOTES

1. Marie Sander, 'Passing Through Shanghai: Ethnographic Insights into the Mobile Lives of Expatriate Youths,' http://heiup.uni-heidelberg.de/catalog/book/48
2. Cluster of Excellence 'Asia and Europe in a Global Context: The Dynamics of Transculturality' at Heidelberg University, http://www.asia-europe.uni-heidelberg.de/en/

3. Andrea Hacker, 'Building It Together: Collaboration in University-Based Open Access Book Publishing,' *Insights* 27, S (2014): 26–29
4. Transcultural Studies, http://www.transculturalstudies.org and Transcultural Research — Heidelberg Studies on Asia and Europe in a Global Context, http://www.springer.com/series/8753
5. All articles on eLife can be viewed with Lens. See, for example, http://lens.elifesciences.org/14226/index.html (accessed 6 July 2016). The viewer is open source, and its code is available at github, available at https://github.com/elifesciences/lens (accessed 6 July 2016).
6. The statistics for the first two books published by heiUP show that the decision to offer HTML in addition to PDF and PoD was justified. For example, the viewing figures for heiUP's first German book show that since its publication in July 2015 (until the time of this writing), the HTML of the book was accessed three times as often as the PDF (1118 versus 228), while the downloads are the opposite (530 versus 1505), http://heiup.uni-heidelberg.de/catalog/book/43 (accessed 8 July 2016).
7. Hacker, 'Building It Together,' 28
8. This experience tallies with findings of the recent *UK Survey of Academics 2015* that show that when deciding where to publish their research, scholars increasingly check whether 'the journal makes its articles freely available on the Internet, so there is no cost to purchase or read.' In fact, this consideration is up ten percentage points in importance, from 29 per cent in 2012 to 39 per cent in 2015. See Jessica Gardner, 'The Impact of Open Access Mandates. Looking at Trends in the UK Survey of Academics,' *Ithaka S+R Blog*, available at http://www.sr.ithaka.org/blog/the-impact-of-open-access-mandates/.
9. Heidelberg Studies on Transculturality, http://heiup.uni-heidelberg.de/series/info/hst
10. meTypeset, available at https://github.com/MartinPaulEve/meTypeset, was built by Martin Eve as an open-source tool 'to convert from Microsoft Word.docx format to NLM/JATS-XML for scholarly/scientific article typesetting.'
11. An additional complication that came with this workaround of producing PDFs with InDesign lay in the fact that the discrepancies between the necessary functions of MS Word and OpenOffice templates that were compatible with our typesetting program and those required by InDesign were so great that significant time was lost in the process of creating a template that was also compatible with InDesign.
12. This last step in the process required running the XML through meTypeset to check for correctness in terms of parsing, formatting issues, completeness, and accuracy of the bibliography, and to link the coding of the bibliographic entries

to the location of their in-text citations for cross-referencing within the HTML. The program is such that progress cannot be saved during this process, so the book's introduction, which contained more references than any other part, had to be run through meTypeset at least five times (a count was not maintained after that). Each time a glitch or inconsistency was discovered, the process had to start all over. Redoing could result from a missing or incomplete reference, or from a reference accidentally linking to the wrong citation, a common mistake due to the program's inability to show exactly which reference required linking if more than one similar option was available. After finishing the book's introduction and getting a sense of what types of problems could be expected, we made a few modifications to the format of the bibliography file. Part I then required three checks. Parts II, IV, and V only needed two. We encountered further issues with Part III, however, since a table (the book only had one) was causing a significant glitch in the HTML viewer. That part had to be processed at least four times.